

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 1, 2003, 08:50:46 ; Search time 45.5 Seconds
(without alignment)
5869.482 Million cell updates/sec

Title: US-09-768-781-2

Perfect score: 2543

Sequence: 1 atgaacacagaccaca.....caaggcaagtgtgtctga 1309

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09768781/runat_01042003_084730_4952/app_query.fasta_1.1543
-DB=PIR_73 -QMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptp -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09768781@cgn_1_174@runat_01042003_084730_4952 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	932	36.6	444	2	McLeod syndrome-as
2	131	5.2	439	2	hypothetical prote
3	121.5	4.8	783	2	hypothetical prote
4	120.5	4.7	382	2	hypothetical prote
5	120.5	4.7	745	2	cytochrome b245 be
6	116.5	4.6	506	2	sopre germination
7	116.5	4.6	2143	2	hypothetical prote
8	115.5	4.5	498	2	NADH2 dehydrogenas
9	114	4.5	508	2	hypothetical prote
10	112	4.4	459	2	NADH2 dehydrogenas
11	111.5	4.4	785	2	hypothetical prote
12	111	4.4	459	2	hypothetical prote
13	111	4.4	459	2	NADH2 dehydrogenas
14	110.5	4.3	459	2	NADH2 dehydrogenas

15	109.5	4.3	461	2	B97305	probable cation ef
16	109.5	4.3	483	2	T20895	hypothetical prote
17	109	4.3	1154	2	T48829	related to SREBP c
18	108.5	4.3	741	2	F90739	probable transport
19	108.5	4.3	741	2	H85589	probable transport
20	108.5	4.3	786	2	H64817	probable membrane
21	108	4.2	740	2	AH0600	probable membrane
22	106.5	4.2	341	2	G82880	ferric anguibactin
23	106.5	4.2	412	2	T34138	hypothetical prote
24	106	4.2	443	2	C70190	K+ transport prote
25	105.5	4.1	339	2	C86874	ABC transporter pe
26	105.5	4.1	503	2	D75104	transmembrane tran
27	105	4.1	459	2	T17181	NADH2 dehydrogenas
28	105	4.1	459	2	T17144	NADH2 dehydrogenas
29	104.5	4.1	493	2	C86565	oligopeptide perme
30	104.5	4.1	493	2	D72059	peptide ABC transp
31	104.5	4.1	521	2	T15322	hypothetical prote
32	103.5	4.1	378	2	T25628	hypothetical prote
33	103.5	4.1	599	2	G90476	probable Na+/H+ an
34	102.5	4.0	460	2	T13881	NADH2 dehydrogenas
35	102	4.0	474	2	S63658	NADH2 dehydrogenas
36	102	4.0	570	2	D86304	hypothetical prote
37	102	4.0	1584	2	T00026	brain-specific ang
38	101.5	4.0	346	2	T19008	hypothetical prote
39	101.5	4.0	354	2	B72688	probable high-affi
40	101.5	4.0	379	2	D96506	hypothetical prote
41	101	4.0	384	2	S37608	NADH2 dehydrogenas
42	100.5	4.0	264	1	S02157	H+-transporting tw
43	100.5	4.0	358	2	B83808	hypothetical prote
44	100.5	4.0	780	2	H84685	probable vacuolar
45	100	3.9	370	2	C69309	conserved hypothet

ALIGNMENTS

RESULT 1

I39294

McLeod syndrome-associated protein XK - human

N/Alternate names: probable membrane transport protein

C/Species: Homo sapiens (man)

C/Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 01-Dec-2000

C/Accession: I39294; S69126

R/Ho, M.; Chelly, J.; Carter, N.; Danek, A.; Crocker, P.; Monaco, A.P.

Cell 77, 869-880, 1994

A/Title: Isolation of the gene for McLeod syndrome that encodes a novel membrane transpo:

A/Reference number: A54300; MUID:94273191; PMID:8004674

A/Accession: I39294

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-444 <RES>

A/Cross-references: EMBL:Z32684; NID:G515872; PID:G515873

R/Khamilichi, S.; Bailly, P.; Blanchard, D.; Goossens, D.; Carton, J.P.; Bertrand, O.

Eur. J. Biochem. 228, 931-934, 1995

A/Title: Purification and partial characterization of the erythrocyte Kx protein defici

A/Reference number: S69126; MUID:95255304; PMID:7737196

A/Accession: S69126

A/Status: preliminary

A/Molecule type: protein

A/Residues: 7-22 <KHA>

C/Genetics:

A/Gene: GDB:XK

A/Cross-references: GDB:120499; OMIM:314850

A/Map position: Xp21.2-Xp21.1

C/Keywords: phosphoprotein; transmembrane protein

Alignment Scores:

Pred. No.: 3.05e-80 Length: 444
Score: 932.00 Matches: 179
Percent Similarity: 64.43% Conservative: 80
Best Local Similarity: 44.53% Mismatches: 133
Query Match: 36.65% Indels: 10
DB: 2 Gaps: 4

US-09-768-781-2 (1-1389) x I39294 (1-444)

Qy	136	TTCCACATTAGACATCTTTCTCCACCTTTTGTGTACTGTGGGAGGCTGCATCTGCTTGT	195
Db	3	PheProAlaSerValLeuAlaLeuValPheLeuPheValAlaGluThrThrAlaLeu	22
Qy	196	TACATGTTAGATCTATCGAAGAATAGTGAACCTTACTCGATCATACACCTTTTCT	255
Db	23	SerLeuSerSerThrTyrArgSerGlyGlyAspArgMetTrpGlnAlaLeuThrLeuLeu	42
Qy	256	TTCTTTATGTTTTCATCCATTATGTGTCAGTTGACCCCTCATTTTGTCCACAGATCTA	315
Db	43	PheSerLeuLeuProCysAlaLeuValGlnLeuThrLeuLeuPheValHisArgAspLeu	62
Qy	316	GCCAAAGATAAACCGCTATCATTTATTTATGCATCTAATCTCTCTTGGAGCCTGTATCAGA	375
Db	63	SerArgAspArgProLeuValLeuLeuHisLeuLeuGlnLeuGlyProLeuPheArg	82
Qy	376	TGTTTGGAGGCCATGATTAAAGTACCTTCACACTGTGGAAGAAAGGAGCAGGAGGAGCCC	435
Db	83	CysPheGluValPheCysIleTyr-----PheGlnSerGlyAsnAsnGluGluPro	99
Qy	436	TATGTCAGCTCCACCCGAAAGAAG---ATGCTAATAGATGGCAGGAGGTCTCATAGAA	492
Db	100	TyrValSerIleThrIysIysArgGlnMetProLysAsnGlyLeuSerGluGluIleGlu	119
Qy	493	TGGAGGTGGGCCACTCCATCCGAGCCCTGGGTATGTCACCGCAATGCTTACAAACGTATG	552
Db	120	LysGluValGlyGlnAlaGluGlyLysLeuIleThrHisArgSerAlaPheSerArgAla	139
Qy	553	TCACAGATCCAAGCCTTCTGGGCTCAGTGGCCGACCTGACCTATCAGCTCTATGTGAGC	612
Db	140	SerValIleGlnAlaPheLeuGlySerAlaProGlnLeuThrLeuGlnLeuTyrIleSer	159
Qy	613	CTGATCTCTGCAGAGTCCCTTGGGTAGAGTGTGCTAAATGGTATTTTCCCTCGTATCT	672
Db	160	ValMetGlnGlnAspValThrValGlyArgSerLeuLeuMetThrIleSerLeuLeuSer	179
Qy	673	GTCACTTATGGGCCACCTTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTAC	732
Db	180	IleValTyrGlyAlaLeuArgCysAsnIleLeuAlaIleValIleLysTyrAspGluTyr	199
Qy	733	AAGATTCGCCTTGGGCCACTAGAAAGTCTCTCTGCATCCACCATCTGGCGGACATTCGAGATC	792
Db	200	GluValLysValAsnArgLeuAlaTyrValCysIlePheLeuThrTrpArgSerPheGluIle	219
Qy	793	ACTTCCGCCTCTCGATTCTGGTGCTCTCTCTCAGCCACTTTGAAATGAAGCGTGTCCCC	852
Db	220	AlaThrArgValValValLeuValLeuPheThrSerValLeuLysThrTrpValVal	239
Qy	853	TTCTAGTGTCAACTTCTCGATCATCTCTTGGAGCCCTGGATTAAAGTTCTGGAGAAGT	912
Db	240	IleIleLeuIleAsnPhePheSerPhePheLeuTyrProTrpIleLeuPheTrpCysSer	259
Qy	913	GGTCCCGAGATGCCCAATAACATTGAGAAAAACTTACGCGGGTCGGCACTCTGGTGGTC	972
Db	260	GlySerProPheProGluAsnIleGluLysAlaLeuSerArgValGlyThrThrIleVal	279
Qy	973	CTGATTTTCAGTCACCATCTCTATGTCGGCATCACTTCTTGTGTGTGTCAGCTTTGCAG	1032
Db	280	LeuCysPheLeuThrLeuLeuTyrThrGlyIleAsnMetPheCysTrpSerAlaValGln	299
Qy	1033	TTGAGGTTGGCAGACAGAGATCTGTCGACAAAGGCGACAACTGGGGACATATGGGCTG	1092
Db	300	LeuLysIleAspSerProAspLeuIleSerLysSerHisAsnTrpTyrGlnLeuLeuVal	319
Qy	1093	CACATATAGTGTGAGGTTGGTAGAGAATGTGATCAGGTCTTGGTTTAAAGTTCTTGGGA	1152
Db	320	TyrTyrMetIleArgPheIleGluAsnAlaIleLeuLeuLeuLeuLeuTrpTyrLeuPheLys	339
Qy	1153	GTGAAGATGTACTGAATTAATCTGTCAATCTCTGATTCCTGATTCCTCGACCTAATTATGCTTAT	1212
Db	340	ThrAspIleTyrMetTyrValCysAlaProLeuLeuValLeuGlnLeuLeuIleGlyTyr	359

```

Qy 1213 CTGATTTCCATTGGCTCATGCTCTTTCTCTCCAGTACTTGCATCCATTCGGCTCACTC 1217
Db 360 CysThrAlaIleLeuPheMetLeuValPheTrpGlnPhePheHisProCysLysLysLeu 379
Qy 1273 TTCACCCATAATCTAGTAGAC-----TACCCTCCATTGTGTCTGCG-----TGT 1314
Db 380 PheSerSerValSerGluGlyPheGlnArgTrpLeuArgCysPheCysTrpAlaCys 399
Qy 1315 CACCAG 1320
Db 400 ArgGln 401

RESULT 2
T32470
Hypothetical protein F08Fl.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C:Accession: T32470
R:Fulton, B.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid F08Fl.
A:Reference number: Z21174
A:Accession: T32470
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-439 <FUL>
A:Cross-references: EMBL:AF026213; PIDN:AB71305.1; GSPDB:GN00028; CESP:F08Fl
A:Experimental source: strain Bristol N2; clone F08Fl
C:Genetics:
A:Gene: CESP:F08Fl.5
A:Map position: X
A:Introns: 20/1, 72/3; 137/3; 184/2; 263/3; 289/3; 339/3; 369/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F08Fl.5

Alignment Scores:
Pred. No.: 0.000618 Length: 439
Score: 131.00 Matches: 67
Percent Similarity: 37.39% Conservative: 62
Best Local Similarity: 19.42% Mismatches: 98
Query Match: 5.15% Indels: 118
DB: 2 Gaps: 15

US-09-768-781-2 (1-1389) x T32470 (1-439)

```

328	QY	CCGCTATCATTTATTATGCACTTAATCCCTC-----TTGGGACCTGTTTATCAGATGTTTG	381
		:::	
		:::	
126	Db	ProLeuSerIysMetIleValLeuCysIleValLeuMetGlyProLeuPheTrpTyTyr	145
322	QY	GAGGCCATGATTAAAGTACCTCACACTGTGGGCAAAAGACAGGACGAGGACCCCTATGTC	441
146	Db	LysAlaLeuTyTyrGlyTrpMetPheArgLysSerSerAsnGluAsn-----	161
442	QY	AGCCTCACCCGAAAGAAGATGCTAAATAGATGGCGAGGAGTGCTGATAGATGAGGAGGTG	501
162	Db	-----ThrAspGlyGluLys-----	166
502	QY	GGCCATCCATTCGGACCCCTGGGTATGCACCGCAATGCCTACAAACGTATGTCACAG-----	558
167	Db	-----ArgLysCysPheSerIysMetValGluAla	176
559	QY	-----ATCCAGACCTTCCTGGGCTCAGTGCCCGCCAGCTG	591
177	Db	GluArgAspAlaThrLeuLeuArgPhePheGluAlaPheLeuGluSerAlaProGlnLeu	196
591	QY	-----	591
197	Db	IleIleGlnGlySerIleAlaAlaSerTyTyrPheGlnAsnTyTyrGlnThrGlyThrTyTyr	216
592	QY	ACCTATACGCTCTATGAGCGCTGATCTCTGCAGAGGTTCCCTGGGTAGATGTGTGCTA	651
217	Db	ProTyTTrpLeuTyTyrPheGlnAlaAlaSer-----	226


```
QY 427 GAGGAGCCCTATGTACGCTCACCAGAAAGAGATGCTAATAGATGGCGAGGAGTGCTG 486
Db 838 ThrSerProTyrThrSer-----SerValTyr 846
QY 487 ATAGATGGAGGTGGGCGCACTCATCGGACCCCTGCTATGCAACCGAATGCTACAAA 546
Db 847 LeuGlyTrpAlaMetSerSerGlyIleAlaLeuVal----- 858
QY 547 CGTATGTACAGATCCAAAGCTCTCGGGCTCAGTGGCCCGCAGCTG---ACCTATCAGCTC 603
Db 859 -----ValThrAlaIleLeuProIleValSerTrpPheAlaThrTyrArgPhe 874
QY 604 TATCTGAGCCGTGATCTCTGCAGAGGTTCCCTGGGTAGAGTTGTGCTAATGGTATTTC 663
Db 875 SerHisSer-----SerAlaValCysLeuMetIlePheSer 886
QY 664 CTGGTATCTGTACCTATGAGGGCCACCCTTTGGCAATATGTTGGCTATCCAGATCAAGTAC 723
Db 887 ValValLeuValAlaPheCysGlyThr-----SerTyrLeuGluValValLysSerArg 904
QY 724 GAT-----GACTACAGATTCGCTTGGCCCACTAGAGTCTCTC 762
Db 905 AspAspGlnLeuProThrLysGlyAspPheLeuAlaLeuProLeuAlaCysIle 924
QY 763 -----TGCAATCACCACATC 774
Db 925 ProAlaLeuLeuSerLeuCysCysGlyMetValLysTrpLysAspAspCysTrpIleLeu 944
QY 775 TGGCGGACATGGAGATCATCTCCGGCTCTGATCTGGTCTCTTCTCAGCCACTTTG 834
Db 945 SerArgGlyValTyrValPhePheSerIleGlyLeuLeuLeuPheGlyAlaIleAla 964
QY 835 AATTGAGGCTGTG---CCCTTCCTAGT---CTCACTTCCTGATCATCCTTTTGG 888
Db 965 AlaValIleAlaValLysProTrpThrIleGlyValSerPheLeuLeuValLysPhe--- 983
QY 889 CCCTGGATTAACTTCTCGAGAAGTGGTCCAGATGCCAATACATTGAGAAAACTTC 948
Db 983 ----- 983
QY 949 AGCCGGGTGGCACTCTGGTGTCTGATTTCACTACCACTCTCTATGCTGGCATCAAC 1008
Db 984 -----LeuMetValValThrIleGlyValIleHisLeuTrpAlaSerAsnAsn 999
QY 1009 -----TTCTCTGCTGGTCACTTTCAGCTTTCGAGTTGAGTTG 1041
Db 1000 PheTyrLeuThrArgLysGlnThrSerPheValCysPheLeuAlaLeuLeuGlyLeu 1019
QY 1042 GCA-----GACAGAGATCTCGTCGACAAAGGCGAGAACTGGGGACAT 1083
Db 1020 AlaAlaPheLeuLeuGlyTrpHisGlnAspLysAlaPheAlaGlyAlaSerValGlyTyr 1039
QY 1084 ATG-----GGCCTGCACTATATAGTGTGAGGTTGGTAGAAGATGTG 1122
Db 1040 PheThrPheLeuSerLeuLeuAlaGlyArgAlaLeuAlaValSerSerThrGluLeuVal 1059
QY 1123 ATCATGCTCTGTTTAAAGTTCTTTGGAGTGAAGTTACTGAATTAATGTCATCTCC 1182
Db 1060 LeuProValTyrValTyrAsp-----AlaHisAla 1069
QY 1183 TTGATTCCTTGCAGCTCATATTGCTTATCTGATTTCCATTGCTTCATGCTC 1236
Db 1070 AspCysGlyLysAsnValSerAlaAlaPheLeuValLeuTyrGlyIleAlaLeu 1087
```

RESULT 8

```
T11039
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Chlamydomonas eugametos mitochondrion
C:Species: mitochondrion Chlamydomonas eugametos
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11039
R:Denovan-Wright, E.M.; Nedelcu, A.M.; Lee, R.W.
Plant Mol. Biol. 36, 285-295, 1998
```

```
A;Title: Complete sequence of the mitochondrial DNA of Chlamydomonas eugametos.
A;Reference number: Z17244; MUID:98145434; PMID:9484440
A;Accession: T11039
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-498 <DEN>
A;Cross-references: EMBL:AF008237; NID:92865253; PID:92865257; PIDN:AAC19340.1
C;Genetics:
A;Genome: mitochondrion
A;Note: nad4
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Alignment Scores:
Pred. No.: 0.019 Length: 498
Score: 115.50 Matches: 65
Percent Similarity: 37.46% Conservative: 44
Best Local Similarity: 22.34% Mismatches: 89
Query Match: 4.54% Indels: 93
DB: 2 Gaps: 14
```

US-09-768-781-2 (1-1389) x T11039 (1-498)

```
QY 592 ACCATCAGCTCTATGTAGCGCTGATCTCTGCAGAGGTCCCTGGGTAGAGTTGTGCTA 651
Db 40 ThrTyrCysLeuPheValThrMetVal-----ValLeuLeu 51
QY 652 ATGGTATTTCCCTGGGTATCTGTCACTATGGGGCC----- 687
Db 52 TrpThrLeuTyrLeuTrpSerIle---TyrAspAlaIleGlyHisSerLeuGlnMetVal 70
QY 688 ACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAGATTCGCCTT--- 744
Db 71 ValIleIleGluArgLeuHisIleSerPheGlyValAspSerMetSerLeuSerLeuThr 90
QY 745 -----GGGCCCATAGAGTCTCTGCAATCACCATCTGGCGGAGCATTTGGAG----- 789
Db 91 LeuLeuThrSerAlaLeuPheProIleCysIleMetValMetArgThrPheLysGlyTyr 110
QY 790 ATCATCTCCCGCTCTCTGATCTGTGCTCTTCTCAGCCACTTTGAAATTCAGAGCTGTG 849
Db 111 IleThrPheLeuLeuLeuGluIleValIleTyrGlyAlaLeuAsnValLeuAspLeuLeu 130
QY 850 CCCTTC-----CTAGTGCTCAACTCTCTGATCATCTCTCTTTGAG 888
Db 131 GlyPheTyrIleLeuPheGluAlaSerLeuLeuLeuPheLeuLeuIle----- 147
QY 889 CCCTGGATTAAAGTTCTCGAGAAGTGGTCCCGAGATGCCCAATAACATTGAGAAAACTTC 948
Db 148 -----GlyArgSerValTyrGlyAsnIleGluAlaAlaTyr 159
QY 949 AGCGGGTCT-----GGCACTGTGTGTGCTCTGATTTCACTCACCATCTCTC 993
Db 160 LysIleValLeuTyrThrMetGlyGlySerLeuIleLeuLeuProIleIlePheValLeu 179
QY 994 TAGTCTGGC-----ATCAACTTC----- 1011
Db 180 TyrAlaGlnGlySerThrSerLeuIleTyrLeuLeuCysAsnPheGlySerGlnHis 199
QY 1012 ---TCTTGTCTGGTCAGCTTTGCGAGTTGGAGTTGGCAGACAGATCTCTGTCACAAAGGG 1068
Db 200 AlaAspAlaTyrGlnSerAlaTyrTyrLeuLeuSerGluArgGlnMetIle----- 216
QY 1069 CAGAACTGGGACATATGGGCTCGCATATAGTGTGAGTTGGTAGAATGTGTATCATG 1128
Db 217 LeuGlyTrpGly----- 220
QY 1129 GTCTTGTGTTTAAAGTCTTTGGAGTCAAACTGTTACTGAATTACTGATTCCTTGTATT 1188
Db 221 -----LeuPheIleValPheAlaValIleProLeuMetProValHisLeuTrpLeu 238
QY 1189 -----GCCTTGCAGCTCATTTATTCGTTATCTGATT 1218
```

Db 239 ProGluAlaHisValAlaalaProThrAlaGlySerValLeuLeuAlaGlyValLeuLeu 258
QY 1219 TCCATTGCTTCATCTCTTTCTTCAGTACTTCATCTGCATCTGCCTCACTCTTCACC 1278
Db 259 LysLeuGly---GlyLeuGlyPheLeuArgPheMetIleProIleLeuProSerPheThr 277
QY 1279 CATATGTTAGTACTACCTCCATTGTGTCTGC 1311
Db 278 AlaSerIlePheProLeuValCysCysMetCys 288
RESULT 9
D90567
hypothetical protein MYP 4440 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: D90567
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: D90567
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-508 <KUR>
A:Cross-references: GB:AL445566; PID:gl4089858; PIDN:CAC13617.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYP 4440
A:Genetic code: SGC3
Alignment Scores:
Pred. No.: 0.0264 Length: 508
Score: 114.00 Matches: 92
Percent Similarity: 35.59% Conservative: 87
Best Local Similarity: 18.29% Mismatches: 168
Query Match: 4.48% Indels: 156
DB: 2 Gaps: 22
US-09-768-781-2 (1-1399) x D90567 (1-508)
QY 31 ACCTCGACAATGACAGAGTTTATGAAATCTCTGAGGAGCCAAATGTGGATCGGTTTCA 90
Db 48 ThrAspSerPheAspAsnIleTyrAlaAspProLysLysPheGluLeuThrLysValPhe 67
QY 91 TCTCTGGAGGAGATGTCATCCGTGGAGCCACCCCGATTACTTTCCATTAGCATC 150
Db 68 IleLeuGlu----- 70
QY 151 CTTTTCCTCCACCTTTTGTGCTGGGAGGTGCATCTGCTTTGTACATGTT----- 204
Db 71 LeuLeuGlnArgSerLeuHis-----AlaSerPheLeuValTyrPheIleAlaThr 87
QY 205 -----AGATCTATCGAAGAATAGTGAACCTTACTGGATGCATACACCTTTTCT 255
Db 88 AlaPheArgLysIleTyrHisAsnHisArgAsnPheTyrIleTyrIleLeuTrpPheSer 107
QY 256 TTTCTTATGTTTTCATCCATT----- 276
Db 108 IleTyrIlePheSerGlyIlePheIleLeuPheSerArgAsnIleProAsnPheIleThr 127
QY 277 -----ATGTCAGTGGACCTCATTTT----- 300
Db 128 ProAsnGluPheValGluLeuGlyTyrIlePheTyrPheValIleTyrValIlePheLeu 147
QY 301 -----GTCCACAGAGATCTAGCCAAAGATAAACCG 330
Db 148 GluIleLeuTyrSerValTyrTrpLeuLeuTyrAsnArgArgValGluArgAspHisLeu 167
QY 331 CTATCATTTATTCATCTAATCTCTTGGACCTGTTATCATGATGTTTGGAGGCCATG 390
Db 168 GluAlaTyrLysLeuLysIleValPheLeu-----IleValSerMetLeuTyrValLeu 185
QY 391 ATTAGTACCTCACATCTGGGAAGAAGAGGAGGAGGAGCCCTATGTCAGCCTCACC 450

Db 186 Ile-----PheSerSerTrpIleAsnLysSerGln-----IleThr 197
QY 451 CGAAGAAGATGCTAATAGAT----- 471
Db 198 ThrAspAlaLysLeuValGluLeuTrpLeuProPheTyrGlnAsnArgIleGlnAlaPhe 217
QY 472 -----GGCAGGAGGTG-----CTGATAGAATGGAGGTGGCCACTCCATC 513
Db 218 ValGluAlaSerSerAspGluIleValGlyLeuIlePheArgGluIleGlyTyrValIle 237
QY 514 CGACCTGGCTATGACCGCAATGCCATCAACACGTATGTACAGATCCAAAGCTTCTCTG 573
Db 238 AlaTyrLeuIleLeu-----IlePheProPhePheSerLeuLeuLeuTyrSer 254
QY 574 GGCTCAGTGGCCAGCTCACC-----TATCAGCTCTAT 606
Db 255 PheTyrLysProLysArgThrLysValSerLysAsnLysLysAsnLeuPheGlnLeuAsn 274
QY 607 GTGAGCCTGATCTCTGCAGAGGTTCCCTGGGTAGAGTTGTCTAATGTTATTTCCCTG 666
Db 275 ValSerIleIleValLeuLeuSerThrPheLeuAsnTyrTrpIleIleLeuPheLysAsn 294
QY 667 GTATCTGTCACTATGCGGCCACCCTTTTGCATATATGTTGGCTATCCAGATCAAGTACGAT 726
Db 295 ValAsnValSerPhe-----LeuGluIleAsn 303
QY 727 GACTACAAGATTCGGCTTGGGCCACTAGAGTCTCTGCATCACC-----ATCTGG 777
Db 304 AsnLeuPheLeuGlyLeuAsnValValPheLeuLeuValPheThrPheLeuIleLeuTrp 323
QY 778 CGGACATTGGAGATCATTCCCGCCTCTGCTGCTCTCTCTCAGCCACTTTGAAA 837
Db 324 AsnArgLeuArgLeuLysSerSer-----IleTyrPheMetPheSerValSerPhe 340
QY 838 TTGAAGGCTGTGCTCTCTAGTGTCAACTTCCTG-----ATCATCTCTTTGAGCCCTGG 894
Db 341 -----PheIleValSerPheThrLeuIleIlePheHis----- 353
QY 895 ATTAAGTTCTGGAGAAGTGGTCCCGCAGATGCCCAATACATTGAGAAAACCTTCAGCCGG 954
Db 354 IleSerTyrAspArg-----GluAsnLysIleSerMet 364
QY 955 GTCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
Db 365 PheAsnIleLeuIlePheSerPheValLeuThrIleLeuPheLeuPheIleLeuPheSer 384
QY 1015 TGCTGCTCAGCTTTCAGTTGAGTTGGCAGACAGAGATCTCGTCGACAAAGGCGAGAAC 1074
Db 385 AsnTyrGluMetSerThrLeuGlnLysMetSerPheAsnLeu----- 398
QY 1075 TGGGACATATGGGCTGCATATAGTGTGAGTGGTAGAGATGTGATCATGTTCTTG 1134
Db 399 -----LeuLeuThrSerPheSerLeuLeu-----TyrIlePheMetSerVal 412
QY 1135 GTTTTAAAGTTCTTGGAGTGAAGTTTACTGAATTACTGTCTTCTGATGCTGCTG 1194
Db 413 IleHisLysLeuAspLeuValIleIleAlaAsnPheLeuAspPheTyrLeuLeuLeu 432
QY 1195 CAGCTCATTTATGCTTATCTGATTTCCATTGGCTGCTCATCTCTCTTCTTCCAGTACTTG 1254
Db 433 TyrLeuLeuValGlyPheAlaValMetLeuAsnPheValPheValPheSerPheTyr 452
QY 1255 CATCATTCGGCTCATCTTTCACCAT-----AATGTA 1287
Db 453 LysSerLeuSerThrLeuLysThrGluLysLysProTrpThrPheAsnLysLysAsnIle 472
QY 1288 GTAGACTAC 1296
Db 473 PheAspTyr 475
RESULT 10
T17163

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mitochondrion
C:Species: Mitochondrion Pygathrix bieti
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C:Accession: T17163
R:Wang, W.; Foretner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xie, J.
Int.: J. Primatol. 18, 305-320, 1997
A>Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A:Reference number: Z18709

A:Accession: T17163
A:Status: preliminary
A: Molecule type: DNA
A:Residues: 1-459 <NAN>
A:Cross-references: EMDL:U92956; NID:92280437; PID:92280440; PIDN:AAD08823.1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:	
Pred. No.:	0.0402
Score:	12.00
Percent Similarity:	36.54%
Best Local Similarity:	19.23%
Query Match:	4.40%
DB:	2
Length:	459
Matches:	80
Conservative:	72
Mismatches:	142
Indels:	122
Gaps:	19

US-09-768-781-2 (1-1389) x T17163 (1-459)

[illegible]

US-09-768-781-2 (1-1389) x A86299 (1-785)

QY 148 ATCCCTTTCTCCACCTTTTGTAC-----TGTGGGAGGCTGCATCT--- 189

Alignment Scores:		
Pred. No.:	0.0495	Length:
Score:	111.50	Matches:
Percent Similarity:	38.2%	Conservative:
Best Local Similarity:	19.69%	Mismatches:
Query Match:	4.38%	Indels:
DB:	2	Gaps:
		785

RESULT 11

A86299
hypothetical protein F309.18 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C;Accession: A86299

R;Theologus, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White-

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;AUTHORS: HUNTER, J.L.; JENKINS, J.; JOHNSON-HORVON, C.; KLAN, S.; KLAUCKIN, C.A.; LEE, Y.; LITVIN, S.Y.; LIU, Z.X.; LYONS, T.S.; MATHIAS, L.J.

C.A.; LT, U.H.; LT, I.; LTH, A.; LTH, S.A.; LTH, Z.A.; LTHOS, O.S.; MARCET,
PITTS M . DOORNOY E . POWELL D . SAKANO H

RIZZO, M.; ROONEY, I.; ROWLEY, D.; SAKALO, H.
 2. Authors: Galzberg S L.; Schwartz J P.; Shinn P.; Southwick A M.; Sun.

ker M : Wu D : Yu G : Fraser C M : Venter J C : Davis R W
R/Audiot: Salzberg, S.D.; Schwartz, G.K.; Smith, E.; Soderwick, A.H.;

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141: MIID:21016719: PMTD:11130712

A: Accession: A86299

A: Status: preliminary

A:Molecule type: DNA

A:Residues: 1-785 <STO>

A;Cross-references: GB:AE005172; NID:g4966359; PIDN:AAD34690.1; GSPDB:GN001

C;Genetics:

A;Map position: 1

Alignment Scores:

Pred. No.:	0.0495	Length:	785
------------	--------	---------	-----

Score:	111.50	Matches:	89
--------	--------	----------	----

Percent Similarity: 38.27% Conservative: 84

Best Local Similarity: 19.69% Mismatches: 172

Query Match:	4.38%	Indels:	107
--------------	-------	---------	-----

DB:	2	20
Gaps:	2	

US-09-768-781-2 (1-1389) x A86299 (1-785)

100

QY 148 ATCCCTTTCTCCACCCTTTTGTA-----TGTGGGAGGCTGCATCT-- 189

```
Db 30 LeuValPheSerGlnPhePheTyrLeuPheLeuLysProCysGlyGlnAlaGlyProVal 49
      :::::||||| ||| ||| |||||:::
Qy 190 -----GCTTTGACATGGTTAGAACTATCGAAAGAAATAGTGA 228
      :::::||||| ||| ||| |||||:::
Db 50 AlaGlnIleLeuAlaGlyIleValLeuSerLeuLeuThrIleIleArgLysValHisGlu 69
      :::::||||| ||| ||| |||||:::
Qy 229 ACTTACTGATGACATACACCTTTTCTTTTATGTTTTCATCATATGCTCCAGTTG 288
      :::::||||| ||| ||| |||||:::
Db 70 PhePheLeuGlnLysAspSerAlaSerTyrTyrIlePhePheSerPheLeuLeuArgThr 89
      :::::||||| ||| ||| |||||:::
Qy 289 ACCCTCATTTT---GTCCACAGAGATCTAGCCAAAGATAAACCCCTATCATTTATG 345
      :::::||||| ||| ||| |||||:::
Db 90 AlaPheValPheLeuIleGlyLeuIleAspLeuAsp-----PheMet 104
      :::::||||| ||| ||| |||||:::
Qy 346 CATCTAATCTCTTGGGACCTGTTATCATAGATGTTTGGAGGCCATGATTAAGTACCTACA 405
      :::::||||| ||| ||| |||||:::
Db 105 LysArganLeuLysAsnSerIleValIleThrLeuGlySerLeuValIleSerGlyIle 124
      :::::||||| ||| ||| |||||:::
Qy 406 CTGTGGAAAGAGAGGAGGAGGAGGAGCCCTATGTGACGCTCACCCGAAAGAGATGCTA 465
      :::::||||| ||| ||| |||||:::
Db 125 IleTrpLeu-----PropheLeuTrpPheLeuIleArgPheMetGln 138
      :::::||||| ||| ||| |||||:::
Qy 466 ATAGATGGCGAG----- 477
      :::::||||| ||| ||| |||||:::
Db 139 IleLysGlyAspPheLeuThrPheTyrLeuAlaPheLeuIleThrLeuSerAsnThrAla 158
      :::::||||| ||| ||| |||||:::
Qy 478 -----GAGTGCTGATAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 519
      :::::||||| ||| ||| |||||:::
Db 159 AlaProValIleArgSerIleIleAspTrpLysLeuHisThrSerGluIleGlyArg 178
      :::::||||| ||| ||| |||||:::
Qy 520 CTGGCTATGACCCGCAATGCCTACAAACGTATGTACAGATCCAAAGCCTTCTGGGCTCA 579
      :::::||||| ||| ||| |||||:::
Db 179 LeuAlaIleSerCysGlyLeuPheIleGluIleThrAsnIle----- 192
      :::::||||| ||| ||| |||||:::
Qy 580 GTGCCCCAGCTGACCTATGACCTATGACGCTGATCTCTGACAGAGGTTCCCTCGGT 639
      :::::||||| ||| ||| |||||:::
Db 193 -----PheIleTyrThrIleValLeuSerPheIleSerGlyThrMetThrAlaAsp 209
      :::::||||| ||| ||| |||||:::
Qy 640 AGAGTTGTGCTAATGATATTTCCCTGATCTGTCTGACCTATGGGGCCACCTTTGCAAT 699
      :::::||||| ||| ||| |||||:::
Db 210 IlePheIleTyrSerPheAlaThrGlyValIleIleLeuThrAsnArgPheLeuAlaSer 229
      :::::||||| ||| ||| |||||:::
Qy 700 ATGTTGGCTATCCAGATCAAGTACATGACATACAAAGATTGCGCTTGGGCCACTAGAA--- 756
      :::::||||| ||| ||| |||||:::
Db 230 TrpLeuProLysArgAsnProLysGluLysTyrLeuSerLysAlaGluThrLeuAlaPhe 249
      :::::||||| ||| ||| |||||:::
Qy 757 ---GTCTCTGCTGATCACCATCTGGCGGACATTGGAG-----ATCATTCTCCGCTC 804
      :::::||||| ||| ||| |||||:::
Db 250 IleIleLeuIleLeuIleAlaLeuThrIleGluSerSerAsnLeuAsnSerThrLeu 269
      :::::||||| ||| ||| |||||:::
Qy 805 CTGATCTGGTG-----CTCTTCTGAGCCACTTTGAAA-----TTG 840
      :::::||||| ||| ||| |||||:::
Db 270 PheValPheIleIleGlyLeuMetPheProArgGluGlyLysThrArgThrLeuIle 289
      :::::||||| ||| ||| |||||:::
Qy 841 AAGGCTGCGCTCTTCTAGTGTCACTTCCGATCATCTCTTGGAGCCCTGGATTAAG 900
      :::::||||| ||| ||| |||||:::
Db 290 GlnArgLeuSerTyrProIleHisGluPheValLeuProValTyrPheGlyTyrIleGly 309
      :::::||||| ||| ||| |||||:::
Qy 901 TTCTGGAGAGTGGTGGCCAGATGCCCAATACATTTGAGAAAACCTTCAGCCGGTCCGC 960
      :::::||||| ||| ||| |||||:::
Db 310 PheArgPheSer-----ValaenSerLeuThrLysArgHisTyrLeuValLeu 325
      :::::||||| ||| ||| |||||:::
Qy 961 ACTCTGGGCTGCTGATTTAGTCACTACCTCTATGCTGGATCACTTCTCTGCTG--- 1017
      :::::||||| ||| ||| |||||:::
Db 326 GlyMetThrValAlaLeuSerLeuGlyLysLeuLeuGlyValLeuPheAlaCysSer 345
      :::::||||| ||| ||| |||||:::
Qy 1018 -----TGGTCAGCTTTGACGTTGAGGTTGCGACACAGAT 1053
      :::::||||| ||| ||| |||||:::
Db 346 PheLeuLysIleProLysGlnTyrTrpLeuPheLeuSerThrMetLeuSerValLys--- 364
      :::::||||| ||| ||| |||||:::
Qy 1054 CTCGTCGAAAGGCGAGAACTGGGGACATATGGGCTGCTCACTATAGTGTGAGTTGGTA 1113
      :::::||||| ||| ||| |||||:::
```

```
Db 365 -----GlyHisIleGlyLeu-----ValleuLeuAsp 373
      :::::||||| ||| ||| |||||:::
Qy 1114 GAGATGTGATCATGCTGCTGTTTAAAGTTCTTTGGAGTGAAAGTGTACTGAATTAC 1173
      :::::||||| ||| ||| |||||:::
Db 374 SerAsn-----LeuMetTyrLysLysTyrPheThrProValVal----- 386
      :::::||||| ||| ||| |||||:::
Qy 1174 TGTCTATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
      :::::||||| ||| ||| |||||:::
Db 387 ---HisAspMetPheValAlaAlaLeuValIleMetThrLeuLeuSerGlyValIleThr 405
      :::::||||| ||| ||| |||||:::
Qy 1228 TTCATGCTCTCTTTTCTCCAG-----TACTTGGCTCCATTGGCTCCTCTTCACC 1278
      :::::||||| ||| ||| |||||:::
Db 406 SerLeuLeuLeuArgSerGlnGluLysSerPheAlaHisIleLysThrSerLeuGluLeu 425
      :::::||||| ||| ||| |||||:::
Qy 1279 CATATGTAGTAGACCTACCTCCATCTGTCTGCTGT 1314
      :::::||||| ||| ||| |||||:::
Db 426 PheAspThrThrGluGluLeuArgValLeuThrCys 437
      :::::||||| ||| ||| |||||:::
RESULT 12
T17166
NADH2 dehydrogenase (ubiquinone) (BC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mit
C:Species: mitochondrion Pygathrix bieti
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C:Accession: T17166
R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A:Reference number: 218709
A:Accession: T17166
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-459 <MAN>
A:Cross-references: EMBL, U92957; NID:q2290441; PID:q2290444; PIDN:AAD08826.1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation.
Alignment Scores:
Pred. No.: 0.0501 Length: 459
Score: 111.00 Matches: 80
Percent Similarity: 36.54% Conservative: 72
Best Local Similarity: 19.23% Mismatches: 142
Query Match: 4.36% Indels: 122
DB: 2 Gaps: 19
US-09-768-781-2 (1-1389) x T17166 (1-459)
Qy 142 TTATGATCTCTTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTCTTTGTACATG 201
      :::::||||| ||| ||| |||||:::
Db 118 PheTyrIleLeuPheGluThrThrLeuIleProThrLeuIleIleIleThrArgTrpGly 137
      :::::||||| ||| ||| |||||:::
Qy 202 GTTAGAATCTATCGAAGAATAGTGAATCTTACTGATGACATACACCTTTCTTTCTTTT 261
      :::::||||| ||| ||| |||||:::
Db 138 AsnGlnAlaGluArgLeuAsnAlaSerThrTyrPheLeuPheTyrThrLeuThr----- 155
      :::::||||| ||| ||| |||||:::
Qy 262 ATGTTTTCATCATGTTGCTCCAGTTGACCTCATTTTGTCCACAGAGATCTAGCCAAA 321
      :::::||||| ||| ||| |||||:::
Db 156 -----GlySerLeuProLeuLeuIleMetLeuLeuPheValHisAsnAsnLeuGlySer 173
      :::::||||| ||| ||| |||||:::
Qy 322 GATAAACCGCTATCATATTTATGATCTAATCTCTCTTGGGACCTGTTATCATGATTTTG 381
      :::::||||| ||| ||| |||||:::
Db 174 -----LeuAsnIle----- 176
      :::::||||| ||| ||| |||||:::
Qy 382 GAGGCCATGATTAACTACCTCACACTGTGGAAGAAAGAGAGGAGGAGGAGGAGGAGGAG 441
      :::::||||| ||| ||| |||||:::
Db 177 -----ProLeuLeu----- 179
      :::::||||| ||| ||| |||||:::
Qy 442 AGCCTCACCCGAAAGAGATGCTAATAGATGGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 501
      :::::||||| ||| ||| |||||:::
Db 180 ThrLeuThrAlaGlnLysLeu-----ThrThrThrTrp----- 190
      :::::||||| ||| ||| |||||:::
Qy 502 GGCCACTCCATCCGAGCCCTGCTATGACCGCAATGCTCAAAACGATGTCACAGATC 561
      :::::||||| ||| ||| |||||:::
Db 191 SerHisAsnLeuThrTrpLeuAlaCys-----Met 200
      :::::||||| ||| ||| |||||:::
```

```
QY 562 CAAGCGCTTCTGGGCTCAGTGGCCAGCTGACCTATCAGCTCTATGTGAGCCCTGATCTCT 621
Db 201 MetAlaPheMetValLysMetProLeuTyrGlyLeuHisLeuTrpLeuProLysAlaHis 220
QY 622 GCAGAGGTTCCCTG---GGTAGAGTTGCTAATGGTATTTCCCTGGTATCTGCACC 678
Db 221 ValGluAlaProIleAlaGlySerMetValLeuAlaValLeuLeuLysLeuGlyGly 240
QY 679 TATGGG-----GCCACCTTTGCAATATGTGTGCTATCCAGATCAAGTAGCAT 726
Db 241 TyrGlyMetMetArgLeuThrSerIleLeuAsnProLeuThrGluTyrMetAlaTyr--- 259
QY 727 GACTACAAGATTGGCTGGCCACCTAGAGTCTCTGCATCACCTATCCGCGGACATTG 786
Db 260 -----ProPheLeuMetLeuSerLeuTrpGlyMetIle 270
QY 787 GAGTACATCTCCGCTCTCTGATTTCTGGTGCTCTTCACGCCACTTTGAAATTAAGGCT 846
Db 271 MetThrSerSerThrCysLeuArgGlnThrAspLeuLysSerLeuIleAlaTyrSerSer 290
QY 847 GTGCCCTTCTAGTGTCAACTTCCCTGATCATCTCTTTGAG---CCCTGGATTAGTTC 903
Db 291 ValSerHisMetAlaLeuValIleMetAlaSerLeuIleGlnThrProTrp----- 307
QY 904 TGGAGAAGTGGTGGCCAGATGTCCTCAATTAACATTGAGAAAAAACTTCAGCCGGGTCCGCACT 963
Db 308 -----SerPheThrGlyAla 312
QY 964 CTGGTGGTCTCTGATT-----TCAGTCACCATCTCTATGCTGGCATC 1005
Db 313 IleValLeuMetIleAlaHisGlyLeuThrSerSerMetLeuPheCysLeuAlaAsnSer 332
QY 1006 AACTTCTCTGTGTGCTGAGTTGGAGTTGGCAGACAGATCTCTCGACAAA 1065
Db 333 AsnTyrGluArgThrHisSer---ArgIleMetLeuLeuSerArgGlyLeuGlnThrLeu 351
QY 1066 -----GGCGAGAAGTGGGGACATATGGCTGCACTAT 1098
Db 352 LeuProLeuMetAlaPheTrpTrpPheAlaAlaAsnLeuThrAsnLeuAlaLeuProPro 371
QY 1099 AGTGTGAGGTTGGTA---GAGAAATGTGATCATGTGCTTGGTTTAAAGTTCTTTGGAGTG 1155
Db 372 ThrIleAsnLeuIleGlyLeuLeuValMetMetThrSerPheSerTrpSerHisVal 391
QY 1156 AAAGTGTACTGAATTAAGTCTATCTGATTCCTGATTCCTGAGCTCATATTATGCTATCTG 1215
Db 392 ThrIleMetLeuThr-----GlyLeuAsnMetLeuIleThrAlaLeu 405
QY 1216 ATTTCCATTGGCTTCATGCTCTTTTCTTCCAG-----TACTTGCAT 1257
Db 406 TyrSerLeu---TyrMetLeuValThrThrGlnArgGluThrLeuThrSerHisIleIle 424
QY 1258 CCATTGGCGCTCACTCTTACC---CATATGTAGTAGACTACCTCCAT 1302
Db 425 AsnMetLysProSerPheThrArgGluAsnMetLeuMetPheMetHis 440

RESULT 13
T17172
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mit
C:Species: mitochondrion Pygathrix bieti
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C:Accession: T17172
R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi
Int. J. Primatol. 18, 305-320, 1997
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A:Reference number: Z18709
A:Accession: T17172
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <NAN>
A:Cross-references: EMBL:U92959; NID:G2290449; PID:G2290452; PIDN:AND08832.1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
```

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation; (

Alignment Scores:

Pred. No.:	0.0501	Length:	459
Score:	111.00	Matches:	80
Percent Similarity:	36.54%	Conservative:	72
Best Local Similarity:	19.23%	Mismatches:	142
Query Match:	4.36%	Indels:	122
DB:	2	Gaps:	19

US-09-768-781-2 (1-1389) x T17172 (1-459)

```
QY 142 TTAGCATCTTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTTGTACATG 201
Db 118 PheTyrIleLeuPheGluThrThrLeuIleProThrLeuIleIleThrArgTyrGly 137
QY 202 GTTAGAATCTATCGAAGAATAGTGAACCTTACTGATGACATACACCTTTTCTTTT 261
Db 138 AsnGlnAlaGluArgLeuAsnAlaSerThrTyrPheLeuPheTyrThrLeuThr- 155
QY 262 ATGTTTTCATCCATTATGTGCCAGTTGACCCCTCATTTTGTCCACAGAGATCTAGCCAAA 321
Db 156 -----GlySerLeuProLeuLeuIleMetLeuPheValHisAsnLeuGlySer 173
QY 322 GATAAACCCGCTCATATTATTATGTCATTAATCTCTTGGGACCTGTTATCAGATGTTG 381
Db 174 -----LeuAsnIle----- 176
QY 382 GAGGCCATGATTAAAGTACCTCACATGTGGAGAAAAGAGGAGGAGGAGCCCTATGTC 441
Db 177 -----ProLeuLeu 179
QY 442 AGCCTCACCCGAAAGAGATGCTAATAGATGGCAGGAGGTGCTGATAGATGGGAGTG 501
Db 180 ThrLeuThrAlaGlnIleLeu-----ThrThrThrTrp----- 190
QY 502 GGCACCTCCATCCGACCCCTGGCTATGTCACCGCAATGCTACAAACGTATGTACAGATC 561
Db 191 SerHisAsnLeuThrTrpLeuAlaCys-----Met 200
QY 562 CAAGCCTTCTCGGCTCAGTGGCCAGCTGACCTATCAGCTCTATGTGAGCCCTGATCTCT 621
Db 201 MetAlaPheMetValLysMetProLeuTyrGlyLeuHisLeuTrpLeuProLysAlaHis 220
QY 622 GCAGAGTTCCTGCTG---GGTAGAGTTGCTAATGGTATTTCCCTGGTATCTGTGCACC 678
Db 221 ValGluAlaProIleAlaGlySerMetValLeuAlaValLeuLeuLysLeuGlyGly 240
QY 679 TATGGG-----GCCACCTTTGCAATATGTGTGCTATCCAGATCAAGTAGCAT 726
Db 241 TyrGlyMetMetArgLeuThrSerIleLeuAsnProLeuThrGluTyrMetAlaTyr--- 259
QY 727 GACTACAAGATTGGCTGGCCACCTAGAGTCTCTGCATCACCTATCCGCGGACATTG 786
Db 260 -----ProPheLeuMetLeuSerLeuTrpGlyMetIle 270
QY 787 GAGTACATCTCCGCTCTCTGATTTCTGGTGCTCTTCACGCCACTTTGAAATTAAGGCT 846
Db 271 MetThrSerSerThrCysLeuArgGlnThrAspLeuLysSerLeuIleAlaTyrSerSer 290
QY 847 GTGCCCTTCTAGTGTCAACTTCCCTGATCATCTCTTTGAG---CCCTGGATTAGTTC 903
Db 291 ValSerHisMetAlaLeuValIleMetAlaSerLeuIleGlnThrProTrp----- 307
QY 904 TGGAGAAGTGGTGGCCAGATGTCCTCAATTAACATTGAGAAAAAACTTCAGCCGGGTCCGCACT 963
Db 308 -----SerPheThrGlyAla 312
QY 964 CTGGTGGTCTCTGATT-----TCAGTCACCATCTCTATGCTGGCATC 1005
Db 313 IleValLeuMetIleAlaHisGlyLeuThrSerSerMetLeuPheCysLeuAlaAsnSer 332
QY 1006 AACTTCTCTGTGTGCTGAGTTGGAGTTGGCAGACAGATCTCTCGACAAA 1065
```

```

Db 333 AsnTyrGluArgThrHisSer---ArgIleMetLeuLeuSerArgGlyLeuGlnThrLeu 351
QY 1066 -----GGGCAGAACTGGGGACATATATGGCCCTGCATAT 1098
Db 352 LeuProLeuMetAlaPheTrpTrpPheAlaAlaAsnLeuThrAsnLeuAlaLeuProPro 371
QY 1099 AGTGTGAGTTGGTA---GAGAATGTGATCATGTCTTGGTTTAAAGTTCTTTGGAGTG 1155
Db 372 ThrIleAsnLeuIleGlyLeuLeuValMetMetThrSerPheSerTrpSerHisVal 391
QY 1156 AAAGTGTACTGAATTACTGTCTTCCCTTGGCTTGCAGCTCATATATGCTTATCTG 1215
Db 392 ThrIleMetLeuThr-----GlyLeuAsnMetLeuIleThrAlaLeu 405
QY 1216 ATTTCCATTGCTTCAGCTCTCTTTCTTCAG-----TACTTGCAAT 1257
Db 406 TyrSerLeu---TyrMetLeuValThrThrGlnArgGlyThrLeuThrSerHisIlelle 424
QY 1258 CCATTGGCTCACTCTTCACC---CATAATGTAGTACACTACCTCCAT 1302
Db 425 AsnMetLysProSerPheThrArgGluAsnMetLeuMetPheMetHis 440

RESULT 14
T17169
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mit
C:Species: mitochondrion Pygathrix bieti
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C:Accession: T17169
R:Wang, W.; Forester, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi
Int. J. Primatol. 18, 305-320, 1997
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A:Reference number: 218709
A:Accession: T17169
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <WAN>
A:Cross-references: EMBL:U92959; NID:G2290445; PID:G2290448; PIDN:RAD08829.1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Alignment Scores:
Pred. No.: 0.0559 Length: 459
Score: 110.50 Matches: 86
Percent Similarity: 36.45% Conservative: 70
Best Local Similarity: 20.09% Mismatches: 147
Query Match: 4.35% Indels: 126
DB: 2 Gaps: 19

US-09-768-781-2 (1-1389) x T17169 (1-459)
QY 142 TTTAGCATCTTTTCTCCACCTTTTGTACTGTGGGAGGTGCATCTGCTTTGTACATG 201
Db 118 PheTyrIleLeuPheGluThrThrLeuIleProThrLeuIleIleThrArgTrpGly 137
QY 202 GTTAGAATCTATCGAAGAATAGTGAACCTTACTGTGATGACATACACCTTTCTTTCTTT 261
Db 138 AsnGlnAlaGluArgLeuAsnAlaSerThrTyrPheLeuPheTyrThrLeuThr 155
QY 262 ATGTTTTCATCCATTATGTCAGTTCAGCTTCATTTTGTCCACAGAGATCTAGCCAAA 321
Db 156 -----GlySerLeuProLeuIleMetLeuLeuPheValHisAsnAsnLeuGlySer 173
QY 322 GATAAACCGCTATCATTTATGATCTAATCTCTTGGGACCTGTTATCAGATGTTTG 381
Db 174 -----LeuAsnIle----- 176
QY 382 GAGGCCATGATTAAGTACCTCACACTGTGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 441
Db 177 -----ProLeuLeu 179
QY 442 AGCTCACCAGAAAGAGATGCTAATAGATGCGGAGGAGGTGCTGATAGATGGGAGGTG 501

```

```

Db 180 ThrLeuThrAlaGlnLysLeu-----ThrThrThrTrp----- 190
QY 502 GGCACTCCATCCGGACCCCTGGCTATGCGCGCAATGCCTACAAACGTATGTCACAGATC 561
Db 191 SerHisAsnLeuThrTrpLeuAlaCys-----Met 200
QY 562 CAGCCCTTCTGGGCTCAGTGCCCGCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCT 621
Db 201 MetAlaPheMetValLysMetProLeuTyrGlyLeuHisLeuTrpLeuProLysAlaHis 220
QY 622 GCAGAGTTCCTCCCTG---GGTAGAGTTGTGCTAATGATGATTTTCCCTGGTATCTGTCTACC 678
Db 221 ValGluAlaProIleAlaGlySerMetValLeuAlaAlaValLeuLeuLysLeuGlyGly 240
QY 679 TATGGG-----GCCACCTTTGCAATATGTTGGCTATCCAGATCAAGTACGAT 726
Db 241 TyrGlyMetMetArgLeuThrSerIleLeuAsnProLeuThrGluTyrMetAlaTyr--- 259
QY 727 GACTACAGATTCCCTTGGGCCACTAGAGTCTCTGTCATCACCACCTCTGGCGGACATTG 786
Db 260 -----ProPheLeuMetLeuSerLeuTrpGlyMetIle 270
QY 787 GAGATCACTTCCCGCTCTCTGATTCCTGCTCTCTCAGCCACTTTGAAATGAAGGCT 846
Db 271 MetThrSerSerThrCysLeuArgGlnThrAspLeuLysSerLeuIleAlaTyrSerSer 290
QY 847 GTGCCCTTCTAGTGCCTCAACTTCTCTGATCATCTCTTTGAG---CCCTGGATTAAAGTTC 903
Db 291 ValSerHisMetAlaLeuValIleMetAlaSerLeuIleGlnThrProTrp----- 307
QY 904 TGGAGAGTGTGTGCCAGATGCCCAATAACATTGAGAAAACTTCAGCGGGTCCGGCACT 963
Db 308 -----SerPheThrGlyAla 312
QY 964 CTGTGTGCTCTGATTC-----TCAGTCACCATCTCTATGCTGGCATC 1005
Db 313 IleValLeuMetIleAlaHisGlyLeuThrSerSerMetLeuPheCysLeuAlaAsnSer 332
QY 1006 AACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
Db 333 AsnTyrGluArgThrHisSer---ArgIleMetLeuLeuSerArgGlyLeuGlnThrLeu 351
QY 1066 -----GGGCAGAACTGGGCACATATGGCCTCGCACTAT 1098
Db 352 LeuProLeuMetAlaPheTrpTrpPheAlaAlaAsnLeuThrAsnLeuAlaLeuProPro 371
QY 1099 AGTGTGAGTTGGTA---GAGAATGTGATCATGTCTTGGTTTAAAGTTCTTTGGAGTG 1155
Db 372 ThrIleAsnLeuIleGlyLeuLeuValMetMetThrSerPheSerTrpSerHisVal 391
QY 1156 AAAGTGTACTGAATTACTGTCTTCCCTTGGCTTGCAGCTCATATATGCTTATCTG 1215
Db 392 ThrIleMetLeuThr-----GlyLeuAsnMetLeuIleThrAlaLeu 405
QY 1216 ATTTCCATTGCTTCATGCTCTCTTTCTTCAGTA-----CTTGCATCCATTCGGCTC 1268
Db 406 TyrSerLeu---TyrMetLeuValThrThrGlnArgGlyThrLeuThrSerHisIlelle 424
QY 1269 ACTCTTCACCATTAATGTAGTAGA-----CTACCTCCATTG 1304
Db 425 AsnMetAsnProSerPheThrArgGluAsnMetLeuMetPheMetHisMetSerProIle 444
QY 1305 TGCTGTGCTGCTACAGACCCCTCG 1328
Db 445 IleLeuLeuSerLeuAsnProSer 452

RESULT 15
B97305
probable cation efflux pump (multidrug resistance protein) [imported] - Clostridium acet
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97305
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

```

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: B97305

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-461 <KUR>

A;Cross-references: GB:AB001437; PIDN:AAK01229.1; PID:GI5026374; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC3295

Alignment Scores:

Pred. No.:	0.0697	Length:	461
Score:	109.50	Matches:	81
Percent Similarity:	39.00%	Conservative:	75
Best Local Similarity:	20.25%	Mismatches:	137
Query Match:	4.31%	Indels:	107
DB:	2	Gaps:	19

US-09-768-781-2 (1-1389) x B97305 (1-461)

Qy	238	ATGACATACACCTTTCTTCTTTATGTTT-----TCATCCATTATGTCGAGTGCACCTC	294
Db	107	IleSerIleSerPheIleIlePheGlnPheLysSerSerIleLeuTyrSerLeuGly---	125
Qy	295	ATTTTGTCCACAGAGATCTACCCAAAGATAA-----CCGCTATCATTA---TTTATG	345
Db	126	-----AlaSerAspAsnThrTyrProMetAlaValGlnTyrIle	138
Qy	346	CATCTAATCTCTTGGACCTGTTATCAGATGTTTGAGGCGCATGATTAAAGTACCTCACA	405
Db	139	SerIleIleLeuPheGlySerIlePheMetCysLeuAlaSerValMetSerTyrIle---	157
Qy	406	CTGTGGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	465
Db	158	-----MetValSerLeuGlyLysThrLysThrLeu	167
Qy	466	ATAGATGGGGAG-----GAGTGCTGATAGATGGGAGTG-----	501
Db	168	LeuTyrThrAsnIleValGlyValLeuLeuAsnIleIleLeuAsnTyrValLeuValIle	187
Qy	502	-----GGCCACTCCATCCGACCTGGCTGATGACCGCAATGCC	540
Db	188	GlnLeuHisMetGlyIleArgGlySerGlyIleAlaThrVal-----	201
Qy	541	TACAAACGTATGTACAGATCCAAAGCTTCTGCGGCTCAGTGCCCGGAGCTGACCTAT---	597
Db	202	-----LeuSerGlnLeuAlaPheValValAlaLeuValGlnPheAlaTyrMet	218
Qy	598	-----CAGCTCTATGTGAGCCTGATCTCTGCAGAG	627
Db	219	AsnLysLysGlnAsnPheLysPhePheGluAsnIleSerThrAsnIleIleSerGlyAsp	238
Qy	628	GTTCCTCCCTGGGTAGAGTTGCTTAATGTTATTTCTCCCTGGTATCTGTCACCTATGGGCC	687
Db	239	Ile-----IleArgGluIleValLeuValGlyPheSerThrPheIleIleGluIleAlaAsp	257
Qy	688	ACCTTTTGCAATATGTTGGCTATCCAGATCAAGTAC-----GATGACTACAAG	735
Db	258	AlaValValSerAlaValLeuAsnAsnValLeuTyrAlaGlyGlyAspSerAlaIle	277
Qy	736	ATTCCGCTTGGGCCACATAGAAGTCTCTGTCATCACCATCTGGCGGACA---TTGGAGATC	792
Db	278	IleMetLeuGlyValIleThrLysValTyrMetPheMetPheIleThrValIleGlyIle	297
Qy	793	ACTTCCCGCTCTGATTTCTGGTCTCTTCTCA-----	825
Db	298	SerSerGlyMetGlnProIleValGlyTyrAsnPheGlyAlaGlyAsnTyrLysAla	317
Qy	826	-----GCCACTTTGAATTTGAAGGCTGTG-----CCCTTCCTAGTGTCTC	864

Search completed: April 1, 2003, 08:59:15

Job time : 63.5 secs

Db	318	LysAspIleLeuLysPheSerLeuLysThrValIleIleSerAlaPheValTyrVal	337
Qy	865	AACCTTCCTCATCTCTTTGAGCCCTGGATTAACTTCTGGAGAAAGTGTGCCAGATG	924
Db	338	GlyPhe---IleIleTrpAlaGlnProLeuIleGlyPhePheLeuLysAspAlaGlnLeu	356
Qy	925	CCCATATAACATTGAGAAAACCTTCAGCCGGGTCGGCACTCTGGTGGCTCCTGATTCAGTC	984
Db	357	ValSerLysThrValSerAlaPheArgIleValIleSerMetLeuProLeuLeuGly---	375
Qy	985	ACCATCCTCTATGTCGGCATCAACTTCTCTGCTGCTGCTGAGCTTTGCAGTTGAGTTGGCA	1044
Db	376	---IleTyrTyrValAlaIleTyrTyr-----	384
Qy	1045	GACAGAGATCTCGTCGACAAAGGCGCAACTGGGGACATATGGCCTGCACCTATAGTGTG	1104
Db	385	-----GlnAlaIleGlyGluAlaArgIleSerPheIleLeu	396
Qy	1105	AGTTGGTAGAAGATGTGATCATG-----GTCTTGGTTTTTAAGTTCTTT	1149
Db	397	SerIleTyrArgGluLeuIleMetPheIleProMetAlaValIleLeuPheLysIleVal	416
Qy	1150	GGAGTCAAA---GTGTTACTGAATTACTGTCATTCCTTGTATTCGCTTGCAGCTCATTT	1206
Db	417	GlyIleAsnGlyValPheIleAlaTyrProLeuThrAspIleIleValIleLeuThrSer	436
Qy	1207	GCTTATCTCATTTCCATTTGGCTTCATGCTCTCTTCTTCCAGTACTTGCATTCATTCGCG	1266
Db	437	ValTyrPheIleArgArgAlaPheLysGluGlnPheAlaGluGluSerIleProLysArg	456